Applied Research Branch
Strategic Policy
Human Resources Development Canada

Direction générale de la recherche appliquée
Politique stratégique
Développement des ressources humaines Canada

Gender Differences in Mathematics and Science Achievement

Roles of Engagement in Out-of-School Activities and Experimentation in Science Lessons of Eighth Graders

R-99-18E.a

by
Yanhong Zhang

December 1999

The views expressed in Applied Research Branch documents are the authors' and do not necessarily reflect the opinions of Human Resources Development Canada or of the federal government.

Les opinions exprimées dans les documents de la Direction générale de la recherche appliquée sont celles des auteurs et ne reflètent pas nécessairement le point de vue de Développement des ressources humaines Canada ou du gouvernement fédéral.

■

The Research Paper Series includes studies and research conducted under the auspices of the Applied Research Branch of Strategic Policy (SP). Papers published in this series consist of secondary research in the form of background studies and literature reviews that support the research efforts of SP.

La série de documents de recherche comprend des études et des travaux de recherche réalisés sous l'égide de la Direction générale de la recherche appliquée, Politique stratégique (PS). Il s'agit notamment de recherches secondaires sous forme d'études de base et d'analyses documentaires qui appuient les efforts de recherche de PS.

104-02983/

Publication Date/Date de parution – Internet 2000
ISBN: 0-662-28428-3
Cat. No./N° de cat. MP32-29/99-18E

General enquiries regarding the documents published by the Applied Research Branch should be addressed to:

Publications Office
Applied Research Branch
Strategic Policy
Human Resources Development Canada
165 Hôtel de Ville St., 7th Floor
Hull, Quebec, Canada
K1A 0J2

Telephone: (819) 994-3304
Facsimile: (819) 953-9077
E-mail: research@spg.org
http://www.hrdc-drhc.gc.ca/arb/

Si vous avez des questions concernant les documents publiés par la Direction générale de la recherche appliquée, veuillez communiquer avec :

Service des publications
Direction générale de la recherche appliquée
Politique stratégique
Développement des ressources humaines Canada
165, rue Hôtel de Ville, 7e étage
Hull (Québec) Canada
K1A 0J2

Téléphone : (819) 994-3304
Télécopieur : (819) 953-9077
Courrier électronique : research@spg.org
http://www.hrdc-drhc.gc.ca/dgra/

Abstract

Using data from the Third International Mathematics and Science Survey, this study attempts to identify some important explanatory factors for mathematics and science achievement for grade 8 students. While female students perform as well as male students in mathematics, they are at a disadvantage in science. Engagement in information-gathering activities is positively associated with both mathematics and science achievement. Engagement in leisure activities, however, is negatively associated with science achievement. Students reporting more hands-on experience in science lessons also perform better in science. More importantly, engagement in both information-gathering activities and hands-on experience seems to benefit boys more than girls. However, more engagement in information-gathering activities by females tends to reduce the observed gender gap in science, which would otherwise be larger.

This paper is part of a set of two papers titled

Gender Differences in Mathematics and Science Achievement:

- *Roles of Engagement in Out-of-School Activities and Experimentation in Science Lessons of Eighth Graders—R-99-18E.a*
- *Evidence From the 1995 TIMSS Concerning Students in the Final Year of Secondary School—R-99-18E.b*

Résumé

À partir des données de la Troisième étude internationale des mathématiques et des sciences, l'étude de Zhang tente de cerner les principaux facteurs explicatifs du rendement en sciences et en mathématiques des élèves de la huitième année. Si les filles obtiennent d'aussi bons résultats que les garçons en mathématiques, leur rendement est inférieur en sciences. La participation à des activités de collecte d'information est liée positivement au rendement tant en mathématiques qu'en sciences. La participation à des activités récréatives est par contre liée négativement au rendement en sciences. Les étudiants qui déclarent un plus grand nombre d'expériences pratiques en sciences obtiennent également de meilleurs résultats. Qui plus est, le fait de participer aux deux types d'activités : collecte d'information et expériences pratiques, semblent profiter davantage aux garçons qu'aux filles. Toutefois, la participation des filles à des activités de collecte d'information tend à diminuer l'écart constaté entre les deux sexes, en sciences, écart qui serait autrement beaucoup plus important.

Ce document fait partie d'un groupe de deux documents :

Gender Differences in Mathematics and Science Achievement

- *Roles of Engagement in Out-of-School Activities and Experimentation in Science Lessons of Eighth Graders—R-99-18E.a*
- *Evidence From the 1995 TIMSS Concerning Students in the Final Year of Secondary School—R-99-18E.b*

Acknowledgements

The author expresses thanks to several individuals who read various drafts and provided comments. They are Gilles Bérubé, Patrick Bussière, Sylvie Daigneault, Patrice de Broucker, Marc Frenette and Thérèse Laflèche. Pierre Foy and David Robitaille provided valuable insights concerning the Third International Mathematics and Science Study data. The author has also benefited immensely from discussions with colleagues at the Centre for Education Statistics at Statistics Canada, including George Butlin, Jeff Frank, Darren Lauzon, Garth Lipps, Jillian Oderkirk and John Zhao, to name a few.

Table of Contents

1. Introduction 7

2. The Gender Gap in Mathematics and Science Achievement 8

3. Research Questions 12

4. Data and Analysis 13

4.1 Sample 13

4.2 Measures 13

4.3 Method of Analysis 15

5. Data Results 17

5.1 Descriptive Results 17

5.2 Determinants of Mathematics Achievement 19

5.3 Determinants of Science Achievement 25

6. Discussion 32

Appendix A: Indicators Used for Principal Component Analysis 35

Appendix B: Descriptions of HLM Models Used in This Analysis 37

Bibliography 39

1. Introduction

Although the overall gender gap in higher education has been closing in Canada, women continue to be under-represented in many mathematics- and science-related programs. For instance, women made up almost 57% of post-secondary graduates in 1995, up from 51% in 1982. In such traditionally male-dominated fields as law and medicine, 58% and 56% of graduates were women, up from 45% and 41% in 1982. This trend indicates the tremendous improvement in the overall educational opportunities for women in Canadian society. Women's participation in higher education, however, remains low in technical fields such as science and engineering. For example, in 1995, fewer than one in five graduates in physical sciences and engineering was female (Statistics Canada; Wannell & Caron 1994).

Gender gaps in these higher education programs and related occupations can be traced to the learning of mathematics and science in grade schools. Since the learning experiences at the high school level play an important role in students' decisions of further education and future career choice, high school mathematics and science seem to continue to work as a filter that prevents women from gaining access to mathematics- and science-related occupations (Sells 1973). Therefore, understanding gender differences in high school mathematics and science, and factors related to such differences, would enable us to address issues surrounding gender equity in a broader context.

Using data from the Third International Mathematics and Science Survey, this study investigates gender differences in mathematics and science achievement in grade 8. The focus of the study is the extent of gender gaps in these two subjects among eighth grade students, and how engagement in out-of-school activities and classroom experiences are related to such gaps.

The rest of this paper is organised as the follows: Section 2 provides a brief review of previous research on the patterns of gender differences in mathematics and science achievement, as well as theories attempting to explain why such differences, if any, might exist. The role of informal experiences, particularly engagement in out-of-school activities will be highlighted. A number of research questions are developed based on the review of literature, and these questions are stated in Section 3. Section 4 describes the sample of grade 8 students, measures to be used for the analysis, and method of analysis. Section 5 presents the data analysis results. Section 6 concludes by briefly discussing the major findings of this research.

2. The Gender Gap in Mathematics and Science Achievement

With regard to gender gaps in mathematics achievement, two recent meta-analyses of previous studies on this topic have produced congruent conclusions (Friedman 1989; Hyde, Fennema & Lamon, 1990). Gender gaps favouring males in mathematics achievement are relatively small or negligible in early school years, but increase at higher grades. In general, gender gaps tend to be larger among the more able students than among students of average ability. Males also tend to have an advantage over females on test items that measure problem-solving skills.[1] Overall, gender gaps have been closing over time.

The issue of gender gap in science achievement has received less attention from researchers than that in mathematics achievement. Two quantitative syntheses (Steinkamp & Maehr, 1983, 1984) found gender gaps favouring males across science domains and schooling levels. Gender gaps were larger in physics than in general science. They were also larger among junior high school students than among those in elementary or high school. Re-analysing Steinkamp and Maehr's results, Becker (1989) confirmed that gender differences vary by subject domain: males have the biggest advantage in physics, followed by general science. The male advantage is relatively small in biology. Using new methodological advances in meta-analysis, Becker found no change in gender gap across schooling levels. The findings of varying gender gaps across science domains are echoed by Lee and Burkam (1996), who document a large male advantage in physical science, but a modest female advantage in life science.

Two distinctively different explanations underlie studies on gender differences in mathematics and science academic achievement: biological and environmental. The biological argument attributes gender differences in mathematics and science to innate cognitive ability. For instance, Benbow and Stanley (1980, 1983) believe that females do not perform as well as males in mathematics because females have inferior innate ability in spatial visualisation. This view is challenged by the socialization theory, according to which gender differences in academic

[1] Willms & Jacobsen (1990) characterise the *growth* in mathematics skills as "fan-spread"—variation in scores increase over time, but students tend to maintain their initial position in the distribution. Most studies reviewed in the meta-analyses use cross-sectional data and therefore examine the *status* of mathematics skills. The patterns of gender differences as described in these studies, however, are congruent to the "fan-spread" characterisation.

outcomes are due to influences of the environment--home, school and community. A comprehensive review of research on gender differences in spatial ability (Linn & Petersen, 1986) found that no conclusive evidence existed to support either argument. On the other hand, overwhelming evidence indicates that gender differences in mathematics and science achievement are decreasing over time (Frazier-Kouassi et al. 1992; Oakes, 1990; Fennema 1985), which is consistent with the socialization theory.

Robitaille and Garden (1996) provide a conceptual framework that attributes student learning outcomes to a combination of individual, classroom, school and system characteristics. In combination, these characteristics constitute "educational opportunities." Such opportunities consist of formal classroom experiences as well as informal, out-of-classroom activities, and are hypothesised to determine learning outcomes. For instance, Nussbaum, Hamilton and Snow (1997) find a strong relationship between achievement and hours spent watching television. In this case, time spent watching television, an out-of-school activity, is considered as an indication of time allocation between leisure and academic pursuits.

Television programs vary in contents. Some have an educational purpose: to help viewers learn concepts or skills similar to those taught at school. Other programs are purely for entertainment. Watching different types of television programs may lead to different experiences for the viewer. Thus it is worthwhile to examine the effects of watching various types of television programs on academic achievement.

Linn and Hyde (1989) suggest that informal experiences are as important as formal learning experiences to mathematics and science achievement. This view is shared by Erickson and Farkas (1987), who find that when choosing answers to science assessment questions, males are more likely to recall experiences outside of formal science classes, whereas females tend to recall incidents that occurred within science lessons.

Females, however, do not have the same kind of informal learning experiences as males. Analysing data from the 1978 National Assessment of Educational Progress survey, Kahle and Lakes (1983) find that males had many more opportunities to experience science-oriented activities than females. Although girls were interested in being more involved in science activities, they clearly lacked the experience. Many more males than females said that they had

visited factories, weather stations and electric plants. Many more males said they had read science articles, watched science shows on TV, or worked on science projects. Astin (1974) finds that boys are more likely than girls to have informal, mathematically related experiences such as playing with scientific toys, participating in mathematical games, and reading mathematics books. Boys also have more problem-solving experiences outside the mathematics classroom (Kimball 1989). Linn and Hyde (1989) note that males have substantially more informal experience with physical science than females do. In addition, elementary and middle-school girls report fewer science experiences than boys (Jones et al. 1992; Jones & Wheatley 1990; Mullis & Jenkins 1988).

Females are also likely to have different learning opportunities in formal instruction. For instance, Eylon and Linn (1989) suggest that because males take more science courses than females, they perform better. This explanation, however, is challenged by Erickson and Erickson (1984), who find no differences in the number of science courses taken by females and males. Teachers are more likely to assign high-ability boys to top mathematics groups than high-ability girls (Hallinan & Sorenson 1987), and tend to interact more frequently with boys than girls in high school mathematics and science classrooms, particularly in groups of high-ability students (Becker 1981; Eccles, et al. 1986; Jones & Wheatley 1990; Oakes 1990; Parsons et al. 1982; Sadker & Sadker 1986, 1994).

In addition to informal and formal experiences, academic self-confidence is also believed to strongly influence achievement. Female students tend to feel less secure when approaching novel or difficult tasks. Several studies have shown evidence that while females have higher course grades and perform better in familiar contexts, male students tend to perform better in novel situations and on competitive tests (Haggerty, 1991; Maccoby & Jacklin 1974; Peterson & Fennema, 1985; Walberg, 1969; Zarega et al. 1986). In a meta-analysis of gender differences in attitudes about mathematics, Hyde et al. (1990b) report that the magnitude of these differences diminished over the previous two decades. However, males still report more encouragement from their parents and teachers than do their female peers. In addition, males stereotype mathematics as a "masculine" domain much more than females, though neither gender holds strongly stereotypical views.

Oakes (1990) conjectures that the effects of individual, environmental and affective factors on mathematics and science achievement may vary by gender. Testing this conjecture, Smith (1992) finds complex interaction effects between gender on the one hand, and age and family patterns on the other, on the science achievement of seventh and ninth graders in the US. Lee and Burkam (1996) find that while males and females gain the same amount of experience in science laboratories, such experiences improve females' science achievement more than males'.

3. Research Questions

Drawing on conjectures developed in previous research with regard to the potential effects of informal and formal experiences, this study explores the potential relationship between student engagement in out-of-school activities and mathematics and science achievement. More specifically, this study intends to explore the following three research questions:

1. Are there gender differences in mathematics and science achievement in grade 8?

2. Is the level of students' engagement in different types of out-of-school activities associated with their mathematics and science achievement? If yes, are these relationships different for female students than for male students?

3. Do students who have more hands-on experiences show higher science achievement? If yes, do the effects of such hands-on experiences vary by gender?

4. Data and Analysis

4.1 Sample

The sample used for this analysis is the cohort of grade 8 Canadian students who participated in the Third International Mathematics and Science Survey (TIMSS). The TIMSS was conducted in 1995, covering more than forty countries. All Canadian provinces and territories except Prince Edward Island participated in the survey. Five provinces – Alberta, British Columbia, Ontario, Newfoundland and New Brunswick (English-speaking schools only) – were over-sampled to allow for inter-provincial comparisons. The survey adopted a multi-stage sampling strategy in that it first sampled schools, then sampled students (in most cases, one entire class) within each school. The data come from four sources: 1) a written test measuring mathematics and science skills; 2) a student survey of home demographic conditions, engagement in out-of-school activities, attitudes towards mathematics and science, and parental (mother's) attitudes relevant to education; 3) surveys of mathematics and science teachers of the sampled students soliciting information on teacher characteristics and instructional practices; 4) a survey of the principals of the schools attended by the sampled students.

The final sample was restricted to those students for whom data on selected individual, teacher (either science or mathematics), and school characteristics are available. Applying such restrictions resulted in a sample of 8,362 students from 384 classes for the study of mathematics achievement, and 8,318 students from 392 classes for the study of science achievement. TIMSS also oversampled certain types of schools. Thus the TIMSS design weights (for students) were used for all analyses.

4.2 Measures

Outcome Variables. The outcome variables in this study are the international scores on written tests in mathematics and science. Such scores are normalised and standardised across students from all participating countries. In mathematics, grade 8 students from Canada scored 527 on average with a standard deviation of 85. In science, their average score was 531 and standard deviation 94. In both subjects, Canadian students performed above the international averages for grade 8 students--513 and 516, respectively (Beaton et al 1996a, 1996b; Robitaille 1996).

Predictors of Mathematics and Science Achievement. Four sets of variables were used as predictors of mathematics and science achievement in this study. The first set describes *basic characteristics of students*. Variables in this set represent student gender, whether students speak the language of the test at home, and immigrant status of parents. A composite of family socio-economic status was constructed using information on parents' educational attainment, as well as the possession of material items in the home.[2] One question in the student survey asks whether students think they usually perform well in mathematics and science. A variable was created based on response to this question to indicate student confidence in each of the two subjects.

Included in this set of variables are also two composites indicating student attitudes towards the two subjects. The student questionnaire collected information on responses to the following questions: to what extent students like the respective subject, think the subject was interesting as well as easy, think the subject was important to everyone's life, and think that in the future they would like to get a job that would involve using knowledge of that subject. Since some of the questions seemed to illicit quite similar information and I believed there was redundancy in the responses, a technique called Principal Component Analysis was employed to reduce the number of variables (Stevens 1986). This technique reduces the number of observed variables into a smaller number of principal components that would account for most of the variance in the observed variables. In this case, two principal components resulted for the process, one representing the attitudes towards mathematics learning and the other science. Some of the variables were reverse-coded so that a higher value in the factor scores indicates more positive attitudes and lower value indicate less positive attitudes.

The second set of variables, the focus of this analysis, are intended to measure *student engagement in various types of out-of-school activities*. The student survey solicited information on the length of time and frequency that students engaged in a variety of out-of-school activities. In addition, students provided information on the frequency with which they watched different types of television programs. Principal component analyses showed these activities could be grouped into three types. The first type is labelled as "knowledge" activities; this includes reading books and magazines, watching television programs on nature, wildlife, and history as well as news and documentaries. The second type, labelled "leisure" activities, includes playing

[2] The author expresses gratitude to Dr. J. Douglas Willms and Mr. George Frempong from the University of New Brunswick for providing an approach to construct the family SES measure.

with friends, engaging in sports, playing computer and video games, and watching television programs that include popular music, sports, cartoons, comedy, adventure or suspense. The third and last type is "cultural" activities, including visiting museums and art exhibitions, attending concerts, and watching television programs on opera, ballet, and classical music. Another variable of interest in this set is an indicator of the frequency with which their school science lessons include hands-on experiments. Again, the variables were coded in such a way that a positive or greater value indicates more engagement and a smaller or less value implies less engagement.

The third set of variables includes *teacher characteristics,* such as the teacher's gender, educational qualifications, and years of teaching experience. The fourth and final set of variables are measures of *school characteristics*. Included in this set are indicators of the type of community (rural, suburban or urban), as well as the provincial jurisdiction where the school is located. Other school characteristics include school size, and proportion of girls enrolled. The data set also has information reported by the principal on the estimated proportions of students in a particular school 1) who came from disadvantaged economic backgrounds; 2) whose parents had no more than primary school education, 3) who came from single-parent families, 4) attended pre-school; 5) whose first language was different from the language taught in the school; 6) had learning, health and/or nutrition problems. Based on such information, a composite statistics was constructed to measure the proportions of disadvantaged students in the school. Greater value of the composite implies that a school has a higher proportion of disadvantaged students. Appendix A provides a more detailed description of the indicators that were used to construct some of the variables in this study.

4.3 Method of Analysis

The major purpose of this analysis was to determine the relationship between mathematics and science achievement of grade 8 students on the one hand, and a number of student, teacher, and school variables on the other. Hierarchical Linear Modeling (HLM) was chosen as the method of analysis to account for the clustering of students within classrooms (Bryk & Raudenbush, 1992). This technique recognises that students within a classroom are not statistically independent observations since they are taught by the same teacher and receive the same instruction. In addition, this technique separates variation among students within a classroom from variation

among classrooms, providing a basis for better understanding why individual students achieve at different levels.

The basic HLM models used in this analysis treat mathematics and science test scores as a function of a set of variables representing student and classroom characteristics. Of particular interest in this analysis were the predictive effects of several student-level variables: gender, their out-of-school activities and, for science achievement, their hands-on experiences in science lessons after taking into consideration other home and school characteristics.

The initial HLM analysis aimed to examine inter- and intra-class variation in grade 8 mathematics and science achievement. Student background variables and indicators representing classroom and school characteristics were added to the models sequentially. Particular attention was paid to the reduction in within-classroom variance and the between-classroom variance as a result of adding these predictors. Finally, the variables of focal interest were entered into the model to detect their relationship to achievement. Potential interactions between gender and other variables were also explored. A more detailed description of the different stages of these models is provided in Appendix B.

5. Data Results

5.1 Descriptive Results

Table 1 displays the means of major variables in the study, comparing females and males. T-tests were conducted for the female-male differences when the variables in question were continuous, and chi-square tests were conducted for the categorical variables. Analysis revealed no significant difference between females and males in mathematics. In science, however, males on average outperformed females by more than one-tenth of a standard deviation.[3] Females reported slightly lower family SES than males, but were quite similar in other demographic characteristics, such as family immigrant status and the frequency of speaking the language of the test at home.

Eighth grade male students showed a consistent advantage in confidence and positive attitudes. Females had a lower self-reported confidence in both mathematics and science. In several aspects of attitudes towards mathematics and science, such as liking the subject, interest, usefulness, level of difficulty, willingness to work in a related area, etc., males generally showed an advantage. The composite index—a measure that is based on these individual indicators—shows that males in general have a more positive attitude towards both mathematics and science than their female peers.

Females and males also differed in the types of out-of-school activities in which they participated. More females reported to read books or magazines frequently, and spent more time reading, than their male peers. A smaller proportion of females said that they often watched TV news and documentaries, as well as programs on nature, wildlife and history. The factor score, however, shows that females were more engaged in "information-gathering" activities than males. Males, on the other hand, reported spending more time on various kinds of leisure and cultural activities, as reflected by higher scores on the respective composite measures. About two-thirds of all students—both males and females—reported that they frequently conducted experiments or practical investigations in science class.

[3] As mentioned earlier, for grade 8 Canadian students, the standard deviation is 85 in mathematics and 94 in science.

Table 1: Means and Standard Errors of Variables Included in the HLM Models
(Standard errors in parentheses)

	Female		Male		Diff. (female-male)	
Mathematics achievement	532.06	(2.48)	528.50	(3.16)	3.56	(3.48)
Science achievement	523.40	(4.37)	537.47	(2.92)	-14.07 ***	(4.63)
Mean family SES	-0.09	(0.042)	0.05	(0.047)	-0.15	(-3.39)
Born out of Canada	0.11		0.13		-0.02	
Speak language of test at home	0.88		0.86		0.02	
Self reported confidence in						
Math (mean=3.18, std dev=0.72)	3.11	(0.03)	3.21	(0.02)	-0.10 ***	(0.03)
Science (mean=0.81, std dev=0.39)	3.03	(0.03)	3.09	(0.03)	-0.06 ***	(0.03)
Composite of Liking Mathematics (mean=0, stddev=0.14)	-0.02	(0.005)	0.00	(0.005)	-0.02 ***	(0.01)
Composite of Liking Science (mean=0, stddev=0.14)	-0.02	(0.007)	0.00	(0.006)	-0.02 ***	(0.01)
Composite indicating level of:						
Cultural activities	-0.02	(0.003)	0.01	(0.004)	-0.04 ***	(0.005)
Leisure activities	-0.05	(0.004)	0.05	(0.004)	-0.10 ***	(0.004)
Information gathering activities	0.01	(0.005)	-0.03	(0.005)	0.03 ***	(0.006)
Percent, experiment in science lesson once a week	0.69		0.67		0.02	
Percent, taught by math teachers who						
Are female	0.41		0.35		0.07 **	
Have degree higher than BA	0.11		0.12		-0.01	
Percent, taught by science teachers who						
Are female	0.41		0.36		0.05 **	
Have degree higher than BA	0.15		0.17		-0.01	
Composite of disadvantaged student population (m=0.04, std=0.83)	0.04	(0.84)	0.04	(0.81)	0.00	(0.04)
School enrolment	643.00	(24.56)	654.00	(21.55)	-10.96	(14.37)
Proportion of female students in school	50.78	(0.94)	48.99	(0.48)	1.78	(0.89)
Rural (%)	14.45		15.22		-0.77	
Urban (%)	49.09		51.91		-2.82	
Suburban (%)	19.07		20.77		-1.70	
Alberta (%)	8.64		8.51		0.13	
British Columbia (%)	10.93		11.36		-0.43	

Table 1 *(continued)*

	Female	Male	Diff. (female-male)
Ontario (English) (%)	28.06	31.60	-3.54
Ontario (French) (%)	1.30	1.24	0.06
Quebec (%)	33.43	29.61	3.62
New Brunswick (%)	1.87	1.75	0.12
Newfoundland (%)	2.37	1.24	1.13
Other provinces (%)	13.40	13.54	-0.14

Notes: 1. Standard deviations are in parentheses.
2. * p<0.10, ** p<0.05, *** p<0.001

A slightly higher proportion of female than male students reported being taught by female mathematics and science teachers. The mathematics teachers of female and male students seemed to have similar educational backgrounds, while small differences emerged in the educational backgrounds of the science teachers of female students and male students. Over eleven percent of the students were taught by mathematics teachers who had graduate degrees, compared to over fifteen percent of students whose science teachers had graduate degrees.

On average, females and males attended schools with similar characteristics in terms of the proportion of disadvantaged students in the population, and the type of community as well as the jurisdictions where the schools were located.

5.2 Determinants of Mathematics Achievement

This section discusses results from the HLM analysis of using mathematics achievement as an outcome variable. No gender differences were detected in mathematics achievement. Engagement in information-gathering and cultural activities was found to be related to scores on the math test. Family SES, confidence in mathematics, language of the test, as well as province of residence were also important predictors of math scores.

The first step of the analysis was to fit an unconditional model—a model without predictors—to partition the total variance in math scores into two components: a student-level component and classroom component. The analysis reveals that student level variation make up 82.1% of the total variation, while classroom-level variation makes up only 17.1% of the total variation in mathematics. This implies that individual differences are much more prominent than differences among classrooms/schools in explaining the variation in mathematics among grade 8 Canadian students.

Table 2: HLM Results for Predicting Mathematics and Science Achievement

	Mathematics		Science	
	Coefficient	SE	Coefficient	SE
Adjusted mean	518.90 ***	(2.24)	518.72 ***	(2.25)
Female student	-1.03	(1.48)	-6.65 **	(2.71)
Engagement in:				
Leisure activities	-6.60	(4.51)	-14.28 **	(7.11)
Cultural activities	-38.59 ***	(4.42)	-44.88 ***	(6.74)
Information gathering activities	8.94 **	(4.39)	24.75 **	(7.05)
Speaking language of test at home	5.22 **	(1.95)	16.28 ***	(2.33)
From immigrant family	-3.06 **	(1.46)	-5.01 **	(1.74)
Family SES	6.65 ***	(0.73)	6.14 ***	(0.86)
Math or science confidence	16.55 ***	(0.90)	12.45 ***	(1.07)
Liking math or science	9.56 **	(4.76)	10.71 **	(5.66)
Teacher characteristics:				
Female	-3.88	(3.81)	1.07	(3.67)
Master's or higher	9.49 *	(5.77)	-5.12	(4.88)
Female student X leisure			26.13 **	(10.46)
Female student X cultural			6.13	(10.16)
Female student X info gathering			-19.06 *	(10.06)
Frequent hands-on experience in science lesson			6.72 **	(2.33)
Female X hands-on experience			-5.97 **	(2.99)
School characteristics				
Disadvantaged student population	-6.43 **	(2.32)	-10.08 ***	(2.10)
School size	-0.01	(0.01)	-0.02 **	(0.01)
Proportion of female students	0.51 *	(-0.27)	0.64 **	(0.30)
Location of school: compared to urban				
Rural	1.08	(5.06)	6.19	(4.94)
Suburban	4.79	(4.85)	9.72 **	(4.49)
Provinces: compared to English Ontario				
Alberta	27.26 ***	(5.59)	49.75 ***	(5.69)
British Columbia	35.17 ***	(7.94)	31.58 ***	(6.49)
French Ontario	-1.19	(8.72)	-43.78 ***	(9.30)
Quebec	62.04 ***	(6.50)	15.03 **	(5.85)
New Brunswick	6.00	(9.34)	7.98	(9.29)
Newfoundland	7.68	(7.70)	2.07	(8.30)
Other	-0.61	(8.05)	25.87 **	(7.64)
Variance reduction				
Between-classroom	50.0 %		47.0 %	
Within-classroom	8.6 %		27.8 %	

Note: * p<0.10, ** p<0.05, *** p<0.001

During the early stages of analysis, potential effects of gender as well as the interactions of gender and other major predictors were explored. No statistically significant relationship between gender and mathematics achievement was detected either before or after a number of variables were entered, representing individual, classroom and school characteristics. Nor did the exploratory analysis reveal any interaction effects between gender and other predictors in the model. Therefore, the results shown in the first column of Table 2 are based on fitting the HLM model that contains a number of variables representing student, family, classroom and school characteristics, including engagement in the three out-of-school activities, but no interaction terms between gender and other variables. The final model explains about 50% of classroom-level variation and 8.6% of student-level variation in mathematics achievement. Since earlier analysis showed that over 80% of the total variation in mathematics scores exist at the student level, the fact that the model explains only 8.6% of the student-level variation implies that the choice of variables representing student characteristics in the TIMSS data needs improvement.

Student characteristics. As mentioned above, no differences between females and males in mathematics achievement were detected. There are, however, strong associations between mathematics and most of the other individual characteristic variables that were entered in the model. Students from families with low SES are at a disadvantage compared to their counterparts from families with high SES. The coefficient on the SES measure is 6.6. We know the difference between the highest-SES and lowest-SES student is 5 on the SES indicator. This implies that the range in mathematics achievement between students from lowest SES families and those from highest SES families is about 33, more than one-third of a standard deviation. Students from immigrant families scored 3 points lower on mathematics test compared to their counterparts whose parents were born in Canada. Students who speak a language at home other than the language of the test scored 5 points lower than those who speak the language of the test at home. Considering that the mean mathematics score for grade 8 was 527 and the standard deviation was 85, these differences seem small.

Mathematics confidence is strongly associated with mathematics achievement. Students who reported usually doing well in mathematics performed better on the test than those who said they did not. The difference in achievement between the most confident students and the least confident students was 65 points, representing almost three quarters of a standard deviation. Students with more positive attitudes towards mathematics also performed better. The difference

was more than one standard deviation between those who had the most positive attitudes overall and those who had the most negative attitudes.

The causal direction between positive attitudes and higher achievement is rarely clear. Psychologists and educators tend to believe that attitudes tend to be an important factor contributing to differences in achievement. It can be argued that achievement differences "cause" differences in attitudes: those students who tend to exhibit more confidence in themselves because they perform better. There is perhaps truth to both propositions. The cross-sectional nature of our data does not allow the examination of to what extent either of these two propositions might be valid in our case.

The types of out-of-school activities that students engage in seem to be related to their mathematics achievement. Two types of activities included in the HLM model--cultural and information-gathering--show a statistically significant relationship with mathematics achievement. There is no evidence, however, that the leisure activities are also related to mathematics achievement. Figures 1 and 2 are graphic presentations of the relationships between mathematics achievement, on the one hand, and information gathering and cultural activities on the other.

Figure 1: Predicted Relationship between Mathematics Achievement and Engagement in Information Gathering Activities

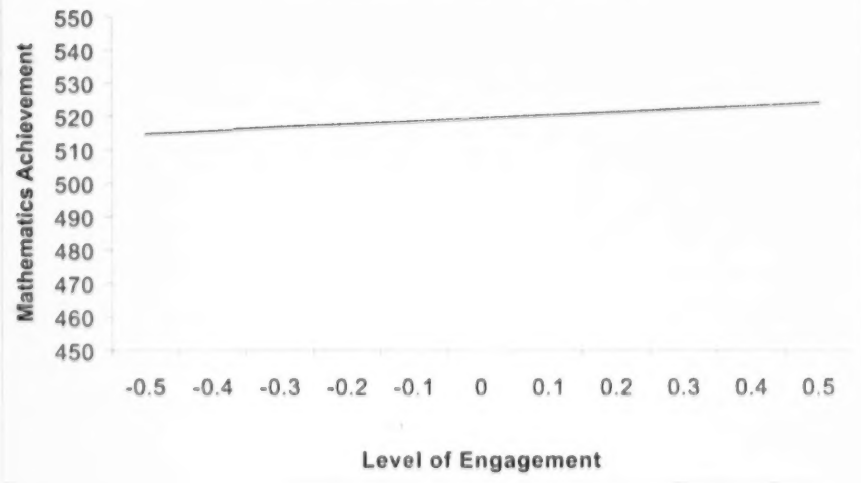
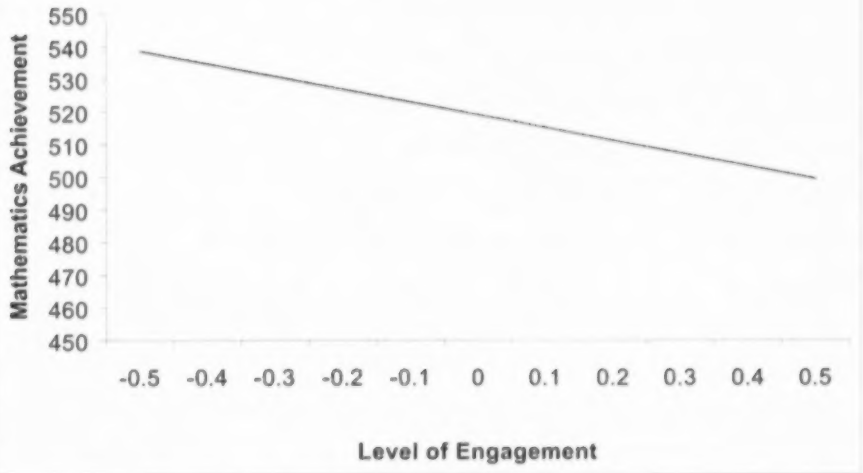


Figure 2: Predicted Relationship between Mathematics Achievement and Engagement in Cultural Activities



In plotting the relationship between achievement and information-gathering activities, all the other variables in the model were set to their mean values. Thus the slope represents the predicted score at each level of engagement for a student who is typical in terms of all other characteristics considered in the model. The positive slope implies that those students who were more engaged in information-gathering activities performed better on the mathematics test. This relationship can be illustrated by two students who share the same characteristics as indicated by the variables in the model, except for their level of engagement in information-gathering activities. Those students who reported the highest level of engagement are predicted to score more than 9 points higher than the one who reported the lowest level of engagement. This difference is equivalent to more than one-tenth of a standard deviation.

The relationship between cultural activities and mathematics achievement was plotted in the same manner. The downward slope indicates a negative relationship between the level of engagement in cultural activities and mathematics achievement. The relatively steep slope implies a strong relationship. To return to the example of the two students who have the same characteristics except for their level of engagement in cultural activities: the student who reported the least engagement in cultural activities is predicted to score 40 points higher than

his/her counterpart who reported the highest level of engagement. Representing almost half of a standard deviation, this difference is relatively large.

Caution should be taken in interpreting the statistical relationships between math scores and engagement in out-of-school activities. The negative association between engagement in cultural activities and test scores does not mean that such activities would necessarily *cause* lower performance. Instead engagement in different types of activities may be an indication of the intrinsic academic orientation of students. For instance, the students who are more engaged in cultural activities may just be very interested in arts and theatre instead of mathematics. Unfortunately, the cross-sectional nature of the data does not allow us to test which of these two explanations would be closer to the reality.

Classroom and School Characteristics. There is no evidence that student mathematics achievement is related to teacher gender. Students taught by a mathematics teacher with a post-graduate degree scored about 9 points—or over one-ninth of a standard deviation--higher than those taught by teachers with an undergraduate degree. This difference is statistically significant at the 10% level.

Higher percentages of disadvantaged students in a school imply lower mathematics achievement. For example, based on the range of value of the variable representing the composition of student population, and the coefficient estimate of the variable, we know that students from schools with the lowest percentage of any combination of disadvantaged groups—either socio-economic or academic—are predicted to score almost 25 points higher than those whose schools have the highest percentage of disadvantaged groups. Students from schools with a higher proportion of female students did slightly better than those from schools with lower proportion of females students, though there was no evidence that this effect was stronger for females than for males. No relationship was detected between mathematics achievement and school size. Students from rural and suburban schools show the same level of math performance as their counterparts from urban schools.

The province of residence seems to be a strong predictor of mathematics achievement. Compared to their counterparts from English-speaking Ontario, students from Quebec, Alberta and British Columbia on average scored higher. There was no evidence that students from English-speaking

Ontario scored either higher or lower than those from French-speaking Ontario, New Brunswick, Newfoundland and the other rest of the participating provinces and territories. The gap between the highest performer—Quebec—and lowest performer—French-speaking Ontario was as large as 63 points, i.e., three quarters of a standard deviation. Recall that these differences exist even after taking into consideration student demographic factors, their daily out-of-school activities, and affective factors, as well as some teacher and school characteristics.

Summary. Student gender does not appear to be related to mathematics performance of grade 8 students. Engagement in specific types of out-of-school activities, however, is. While those who were more engaged in information-gathering activities had higher mathematics scores, those who were more engaged in cultural activities tended to score lower on the mathematics test. The data does not show that the effects of out-of-school activities on mathematics scores vary by student gender.

In addition to engagement in out-of-school activities, important determinants of mathematics test scores are family SES, proportions of disadvantaged student population, and the province of residence. It should also be noted that within-classroom variation is much larger than between-classroom variation.

5.3 Determinants of Science Achievement

This section presents the analysis results for science achievement as an outcome variable. There exists a modest gender difference in science scores, favouring males. Engagement in information-gathering activities was positively associated with scores on the science test, with a stronger relationship for males than for females. Engagement in cultural activities was negatively associated with science scores. While engagement in leisure activities had a negative impact on the science scores of males, its effects were positive for females. In addition, family SES, confidence, language of the test, and province of residence continue to be important predictors of scores in science.

As for achievement in mathematics, analysis of the variance component shows that a much larger proportion of variation in science achievement lies at the student level (86.2%) than at the classroom level (13.8%). The final results of the HLM analysis are presented in Table 2. Exploratory analysis revealed that the effects of interaction between gender, on the one hand, and

out-of-school activities and experimentation in science class, exist in the data, but not interactions between gender and the demographic factors, as well as teacher and school factors. Therefore, the results presented here are based on the HLM model that retains only the interaction terms between student gender and measures of out-of-school activities as well as hands-on science experiments.

Student characteristics. In contrast to the case of mathematics achievement, this research found a gender difference in grade 8 science: female students are disadvantaged. The magnitude of this disadvantage, however, is relatively small: 6.7 points, or less than one-tenth of a standard deviation. As in the case of mathematics achievement, other demographic characteristics are also strongly associated with science achievement. Students from high-SES families scored higher on the test than their peers from low SES families. The coefficient on the SES measure is 6.1. Given that the range of the SES measure is 5, the coefficient of this size indicates that the range in mathematics achievement between students from the families of lowest SES and highest SES is about 30, more than a quarter of a standard deviation. Students from immigrant families are also at some disadvantage: they scored 5 points lower on the science test than their counterparts whose parents were born in Canada. Students who speak a language at home other than the language of the test scored 16 points lower than those who speak the language of the test at home. This difference represents almost one-fifth of a standard deviation.

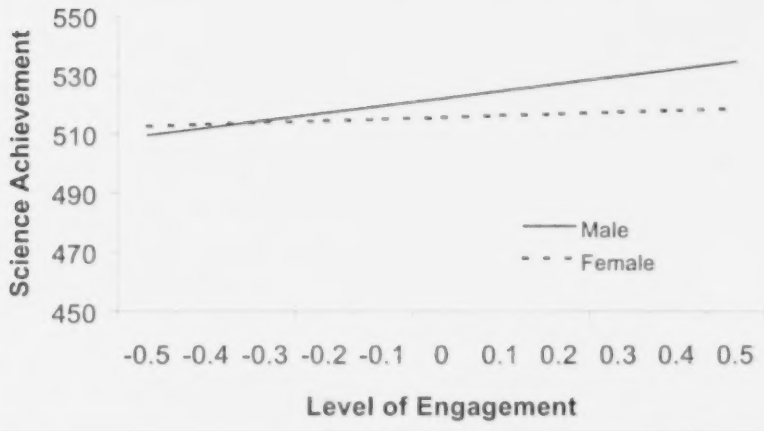
As would be expected, student science confidence is positively associated with science achievement. The difference in scores between the most able/confident students and the least able/confident is 50 points, representing more than half of a standard deviation. Those with more positive attitudes towards science also scored higher. The difference is 11 points, or more than one tenth of a standard deviation, between those with the most positive and most negative attitudes overall.

The relationships between students' engagement in out-of-school activities and their science test scores are more complicated. First, as was true for mathematics achievement, information-gathering activities are positively associated with science achievement. More interesting, the effects are different for the two genders: the positive effects of information-gathering activities are greater for male students than for female students. Second, engagement in cultural activities is negatively associated with science scores, and the effects are the same for both males and

females. Third, engagement in leisure activities is negatively associated with the science scores of males, but positively associated with the scores of females.

Figures 3 to 5 display the relationship between each of the three types of activities and science achievement. On each graph, the solid line represents males and the dotted line represents females. Again, the slopes reflect the change in science achievement associated with each unit of change in the measure of student engagement in a specific type of out-of-school activity, while holding other factors constant. Figure 3 displays the relationship between science achievement and students' engagement in information-gathering activities. The upward slopes indicate that those students who are more engaged in information-gathering activities scored higher on the science test. The effects of such activities on science achievement, however, are greater for males than for females, as indicated by a steeper slope for the line representing males than for the one representing females.

Figure 3: Predicted Relationship between Science Achievement and Engagement in Information Gathering Activities

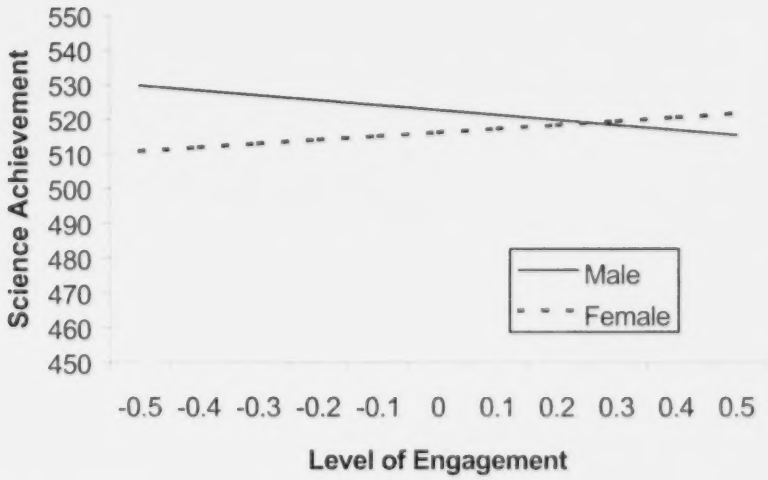


The downward slopes in Figure 4 indicate negative relationships between science achievement and level of engagement in cultural activities. Students who were more engaged in such activities tend to score lower on the test. The effects of cultural activities on achievement are the same for males as for females, as indicated by the equal distance between the two lines at any given level of engagement. Figure 5 shows the relationship between science achievement and leisure

activities. As can be seen, males who are more engaged in leisure activities scored lower on the test. Surprisingly, the opposite holds true for females: those who were more engaged in leisure activities actually scored higher on the science test.

Figure 4: Predicted Relationship between Science Achievement and Engagement in Cultural Activities

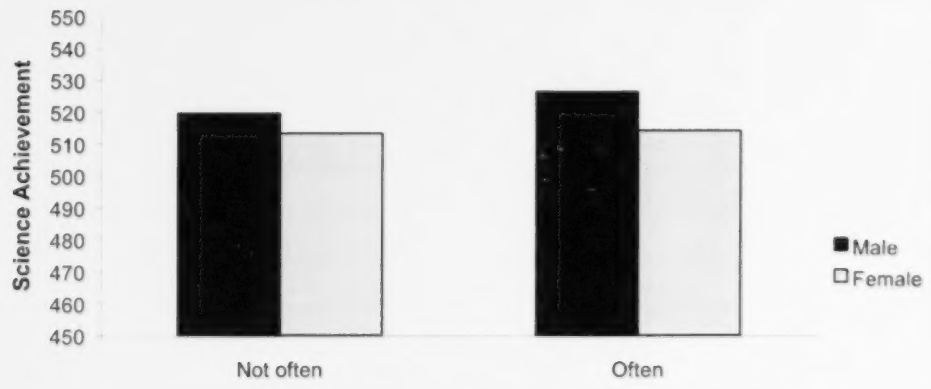
Figure 5: Predicted Relationship between Science Achievement and Engagement in Leisure Activities



Another important finding from this analysis is the complex relationship between hands-on experiments in science lessons and science achievement. In general, students reporting more frequent experiments or practical investigations in their science classes scored higher than their peers who reported less frequent hands-on experiences. The effects of such hands-on experiences, however, seem to be greater for male students than for females.

Figure 6 shows the relationship between the frequency of hands-on science experiences and science achievement, according to gender. As can be seen, males scored higher than females whether or not they reported doing frequent experiments/practical investigations. At the same time, males who reported more frequent hands-on experiences scored higher than those who did not. The same is true for females. However, the gender difference is greater for those reporting more frequent hands-on experiences than for those reporting less frequent hands-on experiences. For instance, in the group with less frequent hands-on experience, males outperformed females by 6 points. In the group with more frequent hands-on experience, the difference is 12 points. This result should in no way be interpreted as implying that hands-on experiments work against female students. Still, it implies that such experiments benefit males more than females.

Figure 6: Predicted Relationship between Science Achievement and Science Experiment



Classroom and School Characteristics. Students taught by female teachers performed as well as those taught by male teachers. There was no evidence indicating that there was any difference

between students taught by teachers with a post-graduate degree and those who were taught by teachers without a post-graduate degree.

Students from schools with higher proportions of disadvantaged students scored lower overall. School size was found to be negatively associated with science achievement, though the effects were quite modest. For instance, students from a school with an enrolment of 500 scored, on average, only 5 points higher than their counterparts from a school of 1,000. Students from schools with a higher proportion of female students also showed slightly higher scores. As was the case with mathematics, students from schools with a higher proportion of females students scored higher in science, though there was no evidence that females from schools with more fellow female students did particularly better. School size was not found to be related to gender differences in achievement.

In terms of location, students from suburban schools on average scored 10 points higher than those from urban schools. Rural school students performed as well as urban school students. As in mathematics, large differences were found in science achievement among students from different jurisdictions. Students from Alberta, British Columbia and Quebec on average performed better than students from English-speaking Ontario. No difference seems to exist between students from New Brunswick and Newfoundland, compared to their counterparts in English-speaking Ontario. Students from French-speaking Ontario on average had the lowest performance. It lags behind the top performing province—Alberta—by almost one standard deviation. Again, it is worthwhile to point out that such differences were detected after controlling from indicators of family SES and a host of other student and classroom characteristics.

Summary. As is the case with mathematics, a much larger proportion of the variation in science achievement lies at the within-classroom than at the between-classroom. The relationship between science achievement and engagement in out-of-school activities is relatively complicated. Engagement in information-gathering activities is positively associated with the test scores of both males and females. Its effects are greater for males than for females. Students who are more engaged in cultural activities tend to score lower on the science test. While males who were more engaged in leisure activities scored lower, the more engaged females tended to score

higher. In addition, family SES, confidence in science, and province of residence continue to be important predictors of scores on the science test for grade 8 students.

6. Discussion

This study examined gender differences in mathematics and science achievement in grade 8, a time when mathematics and science are still compulsory. It focused on the effects of engagement in out-of-school activities on mathematics and science achievement, and the effects of experiments on science achievement. Student demographic and attitudinal as well as classroom and school characteristics were considered. The analysis showed no gender difference in mathematics, but a slight male advantage in science.

Engagement in out-of-school activities was found to be an important predictor of mathematics and science achievement. Several explanations are offered here for the different types of relationships between achievement and engagement in various types of activities. First, engagement in different types of activities leads to quite different learning experiences for individual students. For instance, the educational TV programs re-enforce what children learn in mathematics and science classes, and even may expose the viewer to opportunities of learning new concepts and skills. The same may be happening with the reading of some types of books. Since such "information-gathering" activities provide enriched learning experiences, higher levels of engagement produces high-performing students.

An alternative explanation is that engagement in different types of activities reflects intellectual inclination. Students who are more inclined towards mathematics and science may also tend to read books and magazines, watch both news and documentaries, and television programs on wildlife, nature, and history. This explanation does not necessarily contradicts the first one, though it does not suggest as strong a causal link between time spent on "educational activities" and test scores as the first explanation. Embedded in the second explanation is a two-way relationship: students have different levels of performance because they have different intellectual inclinations, as exhibited through engagement in academic-focused or entertainment-focused activities. Such inclinations, and subsequent devotion—or the lack of—to mathematics and science related experiences lead to varying levels of test performance. On the other hand, it can be argued that the mathematics- and science-inclined students would have higher performance in these tests anyway. In other words, engagement in different types of activities is a manifestation of some kind of inherent ability, which might be a more important determinant of

test scores than the time spent on gaining such educational experiences. Unfortunately, the TIMSS data does not have good measures of ability. As a result, this assertion cannot be tested.

A puzzling finding is that the overall relationship between leisure activities and science achievement is negative for males, but positive for females. With some minor modifications, the explanation of learning experiences and that of ability can be used to explain this relationship. From the perspective of engagement in activities as learning experience, the learning of science (in some domains) involves hands-on experiences. Some of the activities that fall under the category of "leisure," such as playing computer games and sports, may increase such experiences and enhance science achievement. Too much time spent on such activities, however, may distract students from focusing on those activities that are more closely related to academic work. In other words, there may exist a threshold beyond which the activities may work against achievement. Since females are typically less engaged in playing sports and video games, their engagement would tend to stay within such a threshold. On the other hand, males usually spend more time on sports and video games, and more time on such activities may have negative effects on their science achievement.

From the perspective of engagement in activities as a manifestation of inherent ability, one might argue that females who play sports belong to quite different levels of academic and intellectual calibre compared to males who play sports. One might argue that females who play sports tend to be the academically able ones, whereas males who play sports tend to be academically less able. Again, since there is no good measure of student ability, this assertion can not be tested.

It should be noted that even though there is variation along gender lines in science achievement, that variation is relatively small compared to the variation within some other indicators, such as family SES, student confidence, language of the test, and particularly province of residence. Since analysis so far has revealed relatively large inter-provincial differences, future researchers might wish to explore factors that lie underneath such differences. Given that TIMSS data contain rich information on instructional practice and school organisation, a starting point might be to explore the potential relationship between inter-provincial variation in instructional approach and school organisation on the one hand, and academic achievement on the other.

The explanatory "power" of the models used in this analysis seems to be limited since they accounted for very small proportions of the individual-level variation in math and science achievement. The choice of variables is dependent upon the availability of data, and efforts are needed to improve the data in order to come up with a set of predictors that explain a larger proportion of variation in student achievement.

Gender differences in academic achievement should be viewed in a broader perspective. Females continue to be represented in science and engineering fields, even though they have out-numbered males in higher education as a whole. In addition, females also lag behind males in math and science achievement during the final year of secondary school (Lauzon 1999). At the same time, since females usually demonstrate superior performance in internally administered assessments, some argue that male students are at risk. This suggests that gender difference is an evolving issue and new research should explore the specific areas of female disadvantage as well as characteristics of under-performing males.

Appendix A

Indicators Used for Principal Component Analysis

	Female	Male	Difference
Percent of students reporting to:			
Like mathematics	0.71	0.72	-0.01
Enjoy learning math	0.28	0.23	0.05 **
Think math is boring	0.54	0.53	0.01
Think math is an easy subject	0.56	0.45	0.12 ***
Think math is important to everyone's life	0.05	0.05	0.00
Like a job involving using math	0.53	0.43	0.10
Percent of students reporting to:			
Like science	0.64	0.68	-0.04 *
Enjoy learning science	0.31	0.26	0.05 **
Think science is boring	0.56	0.58	-0.02
Think science is an easy subject	0.59	0.51	0.07 ***
Think science is important to everyone's life	0.19	0.21	-0.02
Like a job involving using science	0.56	0.50	0.05 **
Percent of students reporting that mother thinks it is important to:			
Do well in science	0.47	0.51	-0.05 **
Do well in mathematics	0.70	0.70	0.00
Do well in language	0.64	0.64	0.00
Be placed in classes with high-achieving students	0.20	0.27	-0.06 ***
Percent of students spending less than one hour on the following on a normal school day:			
Playing computer games	0.91	0.71	0.20 ***
Doing jobs at home	0.59	0.66	-0.08 ***
Playing or talking with friends outside of school	0.28	0.34	-0.07 **
Reading a book	0.66	0.80	-0.14 ***
Playing sports	0.47	0.28	0.19 ***
Watching TV and videos	0.37	0.19	0.02
Percent of students reporting that they perform the following activities frequently:			
Attend a concert	0.24	0.24	0.00
Go to movies	0.76	0.77	0.00
Visit a museum or art exhibition	0.14	0.18	-0.04 **
Read book or magazine	0.92	0.77	0.14 ***
Go to the theatre	0.24	0.26	-0.02

Appendix A *(Continued)*

Percent of students reporting watching the following kinds of TV programs every day:			
Comedy, adventure, or suspense	0.68	0.70	-0.03
Cartoons	0.29	0.42	-0.13 ***
Nature, wildlife or history	0.04	0.11	-0.07 ***
News or documentaries	0.20	0.28	-0.08 ***
Popular music	0.54	0.51	0.03
Sports	0.30	0.63	-0.33 ***
Video games	0.10	0.41	-0.31 ***

Note: * p<0.10, ** p<0.05, *** p<0.001

Appendix B

Description of HLM Models Used in This Analysis

In this analysis, two-level HLM models were fitted in three stages.[4] The HLM models in each stage are described as the follows.

Stage 1

A two-level model in the first stage is specified as:

$$\text{Level 1: } Y_{ij} = \beta_{0j} + r_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + u_{0j}$$

In the level 1 equation, Y_{ij} is the mathematics or science achievement score of student i in classroom j; B_{0j} is the mean mathematics or science score for classroom j; and r_{ij} is the error term, assumed to be normally distributed with a mean of 0 and variance σ^2. Thus σ^2 represents student-level or within-classroom variance.

In the level 2 equation, γ_{00} represents the grand mean achievement score; and u_{0j} classroom-level error term, assumed to be normally distributed with a mean of zero and variance of τ_{00}. Thus, τ_{00} represents the between-classroom variance in achievement.

Stage 2

In this stage, a number of control variables representing student as well as classroom characteristics were entered in the null model. Thus the student-level model becomes:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(FEMALE)_{ij} + \beta_{2j}(X_1)_{ij} + \beta_{3j}(X_2)_{ij} \ldots + \beta_{kj}(X_{k-1})_{ij} + r_{ij}$$

[4] The design of TIMSS potentially involves clustering at one more level: clustering of classes within schools. However, typically only one class at each grade from one school was included in the sample. This resulted in no within-school variation. As a result, a two-level, instead of three-level, HLM regression was chosen.

where *FEMALE* is a dummy variable indicating student gender (1 if female, 0 if male), X_{k-1} represent (k-1) number of student-level predictors.

The level 2 model in this case becomes:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(W_1)_j + \gamma_{02}(W_2)_j + \ldots + \gamma_{0s}(W_s)_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$$\beta_{2j} = \gamma_{20}$$

....

$$\beta_{kj} = \gamma_{K0}$$

where W_s are a set of variables representing classroom characteristics, such as the gender of the mathematics or science teacher, the teacher's educational qualifications, school size, indicators of resource constraint for the school, the location of the school, etc. u_{0j} represents the variance in adjusted class mean achievement that remains after taking into consideration these level 2 factors. β_{kj} are the pooled within-classroom regression coefficients of Y_{ij} on covariates X_{k-1}.

Note that the effects of gender are also allowed to be varying across classrooms, as indicated by retaining the error term u_{0j}. However, since analysis revealed little variation in the covariate slopes across classrooms, the effects of these covariates are set to be non-varying or fixed.

Stage 3

In the third stage, the level 1 model was expanded to include variables indicating engagement in out-of-school activities and hands-on experience in science lessons. Potential interactions between gender and these variables were also added to the level 1 model. Accordingly the set of level 2 models now include equations that treat the coefficients of these new level 1 variables as outcomes. Since there seemed little variation in these coefficients across classrooms, no error terms are included in the equations.

Bibliography

Astin, Helen. 1974. "Sex differences in mathematical and scientific precocity," in Julian Stanley, Daniel Keating and Lynn Fox (Eds.), *Mathematical Talent: Discovery, Description, and Development*. Baltimore, MD: Johns Hopkins University Press.

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., Smith T. A. 1996a. *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study*. TIMSS International Study Center, Boston College, Chestnut Hill, Massachusetts.

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., Smith T. A. 1996b. *Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study*. TIMSS International Study Center, Boston College, Chestnut Hill, Massachusetts.

Becker, B. J. 1989. "Gender and science achievement: A reanalysis of studies from two meta-analyses." *Journal of Research in Science Teaching*, No. 26, pp. 141-169.

Benbow, C. P. & Stanley, J. C. 1980. "Sex differences in mathematical ability: Fact or artifact?" *Science*. No. 210, pp. 1262-1264.

Benbow, C. P. & Stanley, J. C. 1983. "Sex differences in mathematical reasoning ability: More facts." *Science*. No. 222, pp. 1029-1031.

Eccles, Jacquelynne S. and Jacob, Janis E. 1986. "Social factors shape math attitudes and performance." *Signs: Journal of Women in Culture and Society*. Vol. 11, No. 2, pp. 367-380.

Erickson, G. L. & Erickson, L. J. 1984. "Females and science achievement: Evidence, explanations and implications." *Science Education*. No. 68, pp. 63-89.

Erickson, G. L., & Farkas, S. 1987. "Prior experience: A factor which may contribute to male dominance in science," in J. S. Daniels & J. B. Kahle (Eds.), *Contributions to the Fourth GASAT Conference*. Vol. 1, pp. 49-56. Ann Arbor, MI: University of Michigan.

Eylon, B. & Linn, M. L. 1989. "Learning and instruction: An examination of four research perspectives in science education." *Review of Educational Research*. No. 58, pp. 251-301.

Fennema, E. 1985. "Explaining sex-related differences in mathematics: Theoretical models," in E. Fennema (Ed.), *Educational studies in mathematics*, Vol. 16, pp. 303-320.

Frazier-Kouassi, S, Malanchuk, O., Shure, P., Burkam, D., Gurin, P., Hollenshead, C., Lewis, D., & Soellner-Younce, P. 1992. *Women in mathematics and physics: Inhibitors and enhancers*. Report commissioned by the Alfred P. Sloan Foundation, Center for the Education of Women, and the Departments of Mathematics and Physics, The University of Michigan.

Friedman, L. 1989. "Mathematics and the gender gap: A meta-analysis of recent studies on sex differences in mathematical tasks." *Review of Educational Research*, Vol. 59, pp. 185-213.

Haggerty, S. 1991. "Gender and school science: achievement and participation in Canada." *The Alberta Journal of Educational Research.* Vol. 37, No. 3, pp. 195-208.

Hyde, J. S., Fennema, E., & Lamon, S. J. 1990. "Gender differences in mathematics performance: A meta-analysis." *Psychological Bulletin*, No. 107, pp. 139-155.

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. 1990. "Gender comparisons of mathematics attitudes and affect." *Psychology of Women Quarterly,* No. 14, pp. 299-324.

Jones, M. G. & Wheatley, J. 1990. "Gender differences in teacher-student interactions in science classroom." *Journal of Research in Science Teaching*. Vol. 27, pp 861-874.

Kahle, J. B., & Lakes, M. K. 1983. "The myth of equality in science classrooms." *The Journal of Research in Science Teaching.* Vol. 22, No. 2, pp. 131-140.

Lauzon. D. 1999. *Gender Differences in Science and Math Achievement in the Final Year of Secondary School: Evidence From the 1995 TIMSS.* Ottawa, Ontario: Statistics Canada and Human Resources Development Canada.

Lee, V. E. & Burkam, D. T. 1996. "Gender differences in middle grade science achievement: Subject domain, ability level, and course emphasis." *Science Education.* Vol. 80, No. 6, pp. 613-650.

Linn, M. C., & Hyde, J. S. 1989. "Gender, mathematics, and science." *Educational Researcher.* Vol. 18, No. 8.

Maccoby, E. E. & Jacklin, L. B. 1974. *The Psychology of Sex Differences.* Stanford, CA: Stanford University Press.

Oakes, J. 1990. "Opportunities, achievement, and choice: Women and minority students in science and mathematics," in C. B. Cazden (Ed.) *Review of Research in Education* (Vol. 16). Washington, DC: American Educational Research Association, pp. 153-222.

Parsons, J. E., Kaczala, C. M., & Meece, J. L. 1982. "Socialization of achievement attitudes and beliefs: Classroom influences." *Child Development.* Vol. 52, pp. 322-339.

Peterson, P. L. & Fennema, E. 1985. "Effective teaching, student engagement in classroom activities, and sex-related differences in learning mathematics." *American Educational Research Journal.* Vol. 22, pp 309-335.

Randhawa, B. 1991. "Gender differences in academic achievement: A closer look at mathematics." *The Alberta Journal of Educational Research.* Vol. 37, No. 3, pp. 241-257.

Robitaille, D. F. & Garden, R. A. (ed.) 1996. *Research Questions & Study Design.* TIMSS Monograph No. 2. Vancouver, British Columbia, Canada: Pacific Education Press.

Robitaille, D. F., Taylor, A. R., Orpwood, G. 1996. *TIMSS-Canada Report: Volume 1. Grade 8.* University of British Columbia, Vancouver, British Columbia.

Sadker, M. & Sadker, D. 1986. "Sexism in the classroom: From grade school to graduate school." *Phil Delta Kappan.* Vol. 67, pp. 512-516.

Sadker, M. & Sadker, D. 1994. *Failing at Fairness.* New York, NY: Scribner.

Sells, L. W. 1973. "High school mathematics s the critical filter in the job market," in R. T. Thomas (Ed.), *Developing Opportunities for Minorities in Graduate Education* (pp. 37-39). Berkeley: University of California.

Human Resources Development Canada and Statistics Canada. *The Class of '95: Report of the 1997 National Survey of 1995 Graduates*. Catalogue SP-137-04-99. Internet: http://www.hrdc-drhc.gc.ca/hrdc/corp/stratpol

Steinkamp, M. W., & Maehr, M. L. 1984. "Gender differences in motivational orientations toward achievement in school science: A quantitative synthesis." *American Educational Research Journal,* Vol. 21, pp. 39-59.

Stevens, J. 1986. *Applied Multivariate Statistics for the Social Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Walberg, H. J. 1969. "Physics, femininity, and creativity." *Developmental Psychology.* 1, 47-54.

Wannell, Ted and Caron, Nathalie. 1994. *The Gender Earnings Gap Among Recent Postsecondary Graduates, 1984-92.* Analytical Studies Branch, Research Paper Series No. 68. Ottawa: Statistics Canada.

Willms, J. D. & Jacobsen, S. 1990. "Growth in mathematics skills during the intermediate years: Sex differences and school effects." *International Journal of Educational Research.* No. 14, pp. 157-174.

Zerega, M. R., Haertel, G. D., Tsai, S. & Walberg, H. J. 1986. "Late adolescent sex differences in science learning." *Science Education.* No. 70, pp. 447-460.